

# On the Equilibrium State of a Small System with Random Matrix Coupling to Its Environment

J. L. Lebowitz,

Department of Mathematics and Physics, Rutgers University, USA

L. Pastur

Theoretical Department, Institute for Low Temperatures, Ukraine

## Abstract

We consider a random matrix model of interaction between a small  $n$ -level system,  $S$ , and its environment, a  $N$ -level heat reservoir,  $R$ . The interaction between  $S$  and  $R$  is modeled by a tensor product of a fixed  $n \times n$  matrix and a  $N \times N$  hermitian Gaussian random matrix. We show that under certain "macroscopicity" conditions on  $R$ , the reduced density matrix of the system  $\rho_S = \text{Tr}_R \rho_{SUR}^{(eq)}$ , is given by  $\rho_S^{(c)} \sim \exp\{-\beta H_S\}$ , where  $H_S$  is the Hamiltonian of the isolated system. This holds for all strengths of the interaction and thus gives some justification for using  $\rho_S^{(c)}$  to describe some nano-systems, like biopolymers, in equilibrium with their environment [26]. Our results extend those obtained previously in [16, 17] for a special two-level system.

## 1 Introduction

The properties of a system  $S$ , in contact with a thermal reservoir  $R$ , is an old yet perennial problem of statistical mechanics. Writing the Hamiltonian of the composite system  $S \cup R$  as

$$H_{SUR} = H_S \otimes \mathbf{1}_R + \mathbf{1}_S \otimes H_R + V_{SR} \quad (1.1)$$

where  $H_S$  and  $H_R$  are the Hamiltonians of the system and of the reservoir and  $V_{SR}$  is the interaction between them, the canonical Gibbs density matrix of  $S \cup R$  is given by

$$\rho_{SUR}^{(c)} = \exp\{-\beta H_{SUR}\} / Z_{SUR}. \quad (1.2)$$

The corresponding reduced density matrix of the system is

$$\rho_S = \text{Tr}_R \rho_{SUR}^{(c)} = \exp\{-\beta \tilde{H}_S\} / \tilde{Z}_S. \quad (1.3)$$

Here  $\tilde{H}_S$  is the "effective" Hamiltonian of  $S$ , which will have the form

$$\tilde{H}_S = H_S + \tilde{V}_S \quad (1.4)$$

and  $\tilde{V}_S$  will in general depends on  $\beta$ ,  $H_S$ ,  $H_R$  and  $V_{SR}$  unless  $V_{SR}$  is "negligible" and  $\rho_S$  can be replaced by

$$\rho_S^{(c)} = \exp\{-\beta H_S\} / Z_S. \quad (1.5)$$

The use of (1.5) for the density matrix of  $S$  may be appropriate even when  $V_{SR}$  is not small, if the system  $S$  is macroscopic, e.g., a system in a large box  $\Lambda$ , and the interaction  $V_{SR}$  takes place only along the boundary  $\partial\Lambda$ . Then the calculations of the properties of the system far from the boundary are approximately independent of the interaction  $V_{SR}$ , becoming rigorously so when  $\Lambda \rightarrow \infty$  (there are exceptions when the system is at a first order phase transition). Our concern here is however with the case when  $S$  is small so the above considerations do not apply.

Recent technological advances, making it possible to create and manipulate meso- and nanosystems, including biopolymers, colloidal particles, etc. have brought the problem of a micro-system in contact with a reservoir to the fore. Likewise, in recent studies of the foundation of quantum statistical mechanics [6, 7, 8, 19, 21, 23], where the approach to equilibrium is related to the system-reservoir entanglement the notion of an equilibrium state for non macroscopic systems also plays an important role. In all such cases the nature of the interaction  $V_{SR}$  between the system and the reservoir is clearly important and  $\rho_S$  is not necessarily of the Gibbs form (1.5). A specific example of this importance is the collapse transition in polymers, which depends strongly on the nature of solvent not just on its temperature [10].

We note here that the distinction between the cases (1.3) and (1.5) was made clearly by Jarzynski [12] who showed that his equality between the work done on the system by changing a parameter in  $H_S$  from  $A$  to  $B$  is given by the difference of the free energies  $\tilde{F}_S$  given by  $\log \tilde{Z}_S$  in (1.3) evaluated at the values of the parameter  $A$  and  $B$ . This means that when  $V_{SR}$  is not negligible then this difference will depend on  $H_R$  and  $V_{SR}$  and not just on the temperature  $\beta^{-1}$  of the environment. This distinction is sometimes blurred in the literature both experimental and theoretical where various aspects of this problem are considered. Many of them go under the name of stochastic thermodynamics, where the equilibrium state of a nano-system in contact with an environment at temperature  $\beta^{-1}$  is sometimes implicitly assumed to be described by the distribution (1.5), see, e.g. reviews [13, 24, 26] and references therein. In fact however these studies do not consider the whole microstate of the small system. Instead, one argues that the reservoir degrees of freedom as well as some "internal" degrees of freedom of the small system change very rapidly compared to those observed, which we denote by  $X$ . They can therefore be assumed to be in thermal equilibrium and their effect on the evolution of  $X$  is then just described via some stochastic term (noise). This means in particular that the only effect of the reservoir on the equilibrium state of the  $X$  variables is to specify its temperature, similar to the situation for macroscopic systems. This seems reasonable in some cases (but not in situations like those in the example of the polymer described earlier). One may think then of the interaction term as fluctuating so rapidly that its detailed nature "washes out" as far as its effect on the equilibrium properties, or even the time evolution, of its slow variables are concerned.

In this paper we consider a simple model of such a situation. We do this by making the interaction between the system and reservoir random. We obtain the Gibbs distribution for an arbitrary finite-level quantum system  $S$  using the frameworks of equilibrium statistical mechanics and assuming that the reservoir has certain macroscopic properties.

Note that in this situation we can start, instead of (1.2), which itself requires some justification [12, 13], with the microcanonical distribution for the composite system at a fixed interval  $\Delta$  around a macroscopic energy  $E_{SUR}$ , i.e., with

$$\rho_{SUR}(E_{SUR}) = \chi_{\Delta}(H_{SUR})/\Omega_{SUR}(E_{SUR}), \quad (1.6)$$

where  $\chi_\Delta$  is the indicator of the energy interval

$$\Delta = (E_{S \cup R} - \delta/2, E_{S \cup R} + \delta/2), \quad \delta \ll E_{S \cup R}. \quad (1.7)$$

and

$$\Omega_{S \cup R}(E_{S \cup R}) = \text{Tr}_{S \cup R} \chi_\Delta(H_{S \cup R}). \quad (1.8)$$

We then obtain  $\rho_S$  of (1.4) with the inverse temperature given by [15]

$$\beta = \frac{\partial \mathcal{S}_R}{\partial E_R} \quad (1.9)$$

where

$$\mathcal{S}_R = \log \Omega_R(E_R) \quad (1.10)$$

is the entropy of the reservoir and  $E_R$  is its energy (since for a reservoir which is very large compared to the size of system we have  $S_R \simeq S_{S \cup R}$  and  $E_R \simeq E_{S \cup R}$ ).

The rest of the paper is organized as follows. The model is described in Section 2, the results and discussion are given in Section 3. Sections 4 and 5 contain the proofs. The model is an extension of that introduced and studied in our works [16, 17], where a simple case of a 2-level  $S$  was considered both in the equilibrium and non-equilibrium setting. Here we restrict ourselves to the equilibrium and leave dynamic considerations for a future work.

According to the above, the proofs are given for the microcanonical distribution (1.6) – (1.10) of the composite system. The results for the canonical distribution (1.2) of the composite prove to be a simple corollary of those for the microcanonical distribution and require a version of a standard (in statistical mechanics) saddle point argument.

Note also that large random matrices have been widely used to model a variety of complex quantum systems, including heavy nuclei and atoms, mesoscopic particles, quantum networks, graphs etc., i.e., not necessarily macroscopic, systems [1, 9, 22]. Thus, our setting could model a qubit or a microcluster in a mesoscopic (or even nanoscopic) particle, a quantum dot, small quantum network, etc.

## 2 Model

We describe now the model and quantities that will be considered. Let  $H_R$  be a  $N \times N$  hermitian matrix,  $\{E_j\}_{j=1}^N$  be its eigenvalues and

$$\nu_N(E) = N^{-1} \sum_{j=1}^N \delta(E - E_j), \quad \int_{-\infty}^{\infty} \nu_N(E) dE = 1. \quad (2.1)$$

be its density of states normalized to unity. We assume that  $\nu_N$  converges as  $N \rightarrow \infty$  to a continuous density  $\nu$  in the sense that for any continuous and bounded function  $f$  we have:

$$\lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} f(E) \nu_N(E) dE = \int_{-\infty}^{\infty} f(E) \nu(E) dE, \quad \int_{-\infty}^{\infty} \nu(E) dE = 1. \quad (2.2)$$

Let  $W_R$  be a random hermitian matrix, distributed according to the Gaussian unitary invariant law given by probability density

$$\mathcal{Z}_R^{-1} \exp \left\{ -\text{Tr} W_R^2 / 2 \right\}, \quad (2.3)$$

where  $\mathcal{Z}_R$  is the normalization constant. In other words, we assume that the entries of  $W_R = \{W_{jk}\}_{j,k=1}^N$ ,  $W_{kj} = \overline{W_{jk}}$  are complex and independent for  $1 \leq j \leq k \leq N$  Gaussian random variables such that

$$\mathbf{E}\{W_{jk}\} = \mathbf{E}\{W_{jk}^2\} = 0, \quad \mathbf{E}\{|W_{jk}|^2\} = (1 + \delta_{jk}). \quad (2.4)$$

This is known as the Gaussian Unitary Ensemble (see e.g. [22], Section 1.1).

Let also  $H_S$  and  $\Sigma_S$  be arbitrary  $n \times n$  hermitian matrices. We define the Hamiltonian of our composite system  $S \cup R$  as a random  $nN \times nN$  matrix (cf. (1.1))

$$H_{S \cup R} = H_S \otimes \mathbf{1}_N + \mathbf{1}_n \otimes H_R + \Sigma_S \otimes W_R / N^{1/2}, \quad (2.5)$$

where  $\mathbf{1}_l$  is the unit  $l \times l$  matrix.

It should be noted that our results are valid not only for the special Gaussian distributed interaction (2.3) – (2.4), but for any real symmetric or hermitian random matrix in (2.5), whose entries  $W_{jk}$ ,  $1 \leq j \leq k \leq N$  are independent and satisfy (2.4). However, in this case the techniques are more involved requiring an extension of those of random matrix theory for so called Wigner matrices (see [22], Section 18.3).

In addition, our results are also valid in the case, where  $H_R$  is random and independent of  $W_R$  for all  $N$ . In this case we have to assume that the sequence  $\{H_R\}$  of random matrices is defined for all  $N \rightarrow \infty$  on the same probability space as the sequence  $\{W_R\}$ , is independent of  $\{W_R\}$  and satisfies (2.2) with probability 1.

Having  $H_R$  and  $W_R$  random, we can view the second term in (2.5) as the Hamiltonian of a "typical"  $N$ -level reservoir and the third term as a "typical" interaction between the system and its reservoir. It is worth mentioning that the notion of typicality has recently been used in the studies of the foundations of quantum statistical mechanics, including the form of reduced density matrix in equilibrium [7, 8, 19, 21, 23]. In addition, the randomness (frozen disorder) is a basic ingredient of the theory of disordered systems. Its successful and efficient use is justified by establishing the selfaveraging property of the corresponding results, i.e., their validity for the overwhelming majority of realizations of randomness for macroscopically large systems, see e.g. [18].

In our case the selfaveraging property is valid in the limit  $N \rightarrow \infty$  and is given by Result **I** below. This suggests that in our model the  $N \rightarrow \infty$  limit plays the role of the macroscopic limits in statistical mechanics and condensed matter theory. Note however that in statistical mechanics the density of states of a macroscopic reservoir is multiplicative in its volume and/or in its number of degrees of freedom. This and the macroscopicity of the system lead to the Gibbs form (1.5) of its reduced density matrix, if the system-reservoir interaction is of short range and confined to the boundary of  $S$  [15].

If, however, the system is small, then, as is already noted, its reduced density matrix is not (1.5) in general even if the reservoir is macroscopically large. This is well understood in statistical mechanics and in principle in the stochastic thermodynamics community [12, 13]. This is especially in the frameworks of dynamical approach where the problem was first studied by Bogolyubov [4] for a classical system consisting of a harmonic oscillator interacting linearly with the macroscopic reservoir of harmonic oscillators, and then in a number of interesting and rather general quantum models with macroscopic many-body reservoirs, see, e.g., [2, 3, 5]. Taking into account that according to (2.1) and (2.2) the density of states  $N\nu_N$  of our reservoir is asymptotically additive (but not multiplicative) in  $N$ , we conclude that the Gibbs form of the reduced density matrix seems to be even less likely in our case than

in the spin-boson model and that one needs additional conditions and procedures to obtain the Gibbs distribution (1.5). We will discuss such conditions below. Here we only mention that an asymptotically additive in volume density of states is the case in the one-body approximation of solid state physics.

According to [16] the Gibbs distribution for the two-level ( $n = 2$ ) version of our model can be obtained if we assume that the reservoir consists of large number  $1 \ll J \ll N$  of independent or weakly dependent parts, more precisely, that the normalized density of states  $\nu_N$  of the reservoir has a special "quasi-multiplicative" form (3.17). One can view this assumption as an effort to obtain the multiplicativity of the density of states of the reservoir on a scale which is intermediate (mesoscopic) between the microscopic and macroscopic scales. It is shown below that the same assumption on  $\nu_N$  allows one to obtain the Gibbs distribution (1.5) for an arbitrary finite dimensional  $S$ .

### 3 Results

We are interested in the asymptotic form as  $N \rightarrow \infty$  of the *reduced density matrix*

$$\rho_S^{(N)} = \frac{N^{-1} \text{Tr}_S \chi_\Delta(H_{S \cup R})}{N^{-1} \text{Tr} \chi_\Delta(H_{S \cup R})} \quad (3.1)$$

corresponding to the microcanonical distribution (1.6) – (1.8) of our model composite system (2.5).

Our first result is

*I. There exists a positive definite, trace one, non-random matrix  $\rho_S$  such that for any fixed  $\Delta$  of (1.7) we have with probability 1*

$$\lim_{N \rightarrow \infty} \rho_S^{(N)} = \rho_S. \quad (3.2)$$

The result is proved in Section 4.

To describe the limiting density matrix  $\rho_S$ , we start from the relation

$$\chi_\Delta(H_{S \cup R}) = \mathcal{E}_{H_{S \cup R}}(\Delta), \quad (3.3)$$

where  $\mathcal{E}_{H_{S \cup R}}$  is the resolution of identity (spectral projection) of the hermitian operator  $H_{S \cup R}$  (2.5) corresponding to the spectral interval  $\Delta$  of (1.7). Denoting

$$e_S^{(N)}(d\lambda) = N^{-1} \text{Tr}_R \mathcal{E}_{H_{S \cup R}}(d\lambda), \quad (3.4)$$

we can write, using (1.7), (3.3) and (3.1),

$$\rho_S^{(N)} = \frac{e_S^{(N)}(\Delta)}{\text{Tr}_S e_S^{(N)}(\Delta)}. \quad (3.5)$$

Thus, we have to find the limit

$$e_S(\Delta) = \lim_{N \rightarrow \infty} e_S^{(N)}(\Delta) \quad (3.6)$$

of the  $n \times n$  random hermitian matrix  $e_S^{(N)}(\Delta)$  and then we can write (3.2) as

$$\rho_S = \frac{e_S(\Delta)}{\text{Tr}_S e_S(\Delta)} \quad (3.7)$$

To find  $e_S(\Delta)$  of (3.6) we will use the resolvent

$$G(z) := (H_{S \cup R} - z)^{-1}, \quad \Im z \neq 0 \quad (3.8)$$

of  $H_{S \cup R}$ . Indeed, we have, by the spectral theorem for hermitian matrices,

$$G(z) = \int \frac{\mathcal{E}_{H_{S \cup R}}(d\lambda)}{\lambda - z}, \quad \Im z \neq 0, \quad (3.9)$$

and we write here and below integrals without limits for those over the whole real axis. It follows then from (3.4) and (3.9) that

$$g_S^{(N)}(z) := N^{-1} \text{Tr}_R G(z) = \int \frac{e_S^{(N)}(d\lambda)}{\lambda - z}, \quad \Im z \neq 0. \quad (3.10)$$

Note that  $e_S$  is the matrix valued measure assuming values in  $n \times n$  positive definite matrices and is uniformly bounded in  $N$ , since for any  $\Delta \subset \mathbb{R}$  the matrix  $\mathcal{E}_{H_{S \cup R}}(\Delta)$ , being an orthogonal projection in  $nN$  space  $\mathcal{H}_S \otimes \mathcal{H}_R$ , is of norm one and according to (3.4)

$$\|e_S^{(N)}(\Delta)\| \leq n.$$

Here  $\|A\|$  is the standard matrix norm of a matrix  $A$  and we used the inequality

$$\|\text{Tr}_R A\| \leq nN \|A\|,$$

valid for any hermitian matrix  $A$  in  $\mathcal{H}_S \otimes \mathcal{H}_R$ .

Recall now that for any non-negative finite measure  $m$  on the real axis one can define its *Stieltjes transform*

$$s(z) = \int \frac{m(d\lambda)}{\lambda - z}, \quad \Im z \neq 0, \quad (3.11)$$

which is analytic for  $\Im z \neq 0$  and such that

$$\Im s(z) \Im z > 0, \quad \Im z \neq 0. \quad (3.12)$$

The correspondence between non-negative measures and their Stieltjes transforms is one-to-one, in particular

$$m(\Delta) = \lim_{\delta \rightarrow 0^+} \frac{1}{\pi} \int_{\Delta} \Im s(\lambda + i\delta) d\lambda. \quad (3.13)$$

Besides, the correspondence is continuous with respect to the weak convergence of measures (see e.g. (2.2)) and the uniform convergence of their Stieltjes transforms on a compact set of  $\mathbb{C} \setminus \mathbb{R}$  (see e.g. [22], Proposition 2.1.2).

It is easy to extend (3.11) – (3.13) to the matrix valued positive definite and bounded measures and their matrix valued Stieltjes transforms, whose examples are (3.4) and (3.10) respectively. Thus, to prove (3.7) it suffices to prove that for a compact set of  $\mathbb{C} \setminus \mathbb{R}$  the matrix

valued functions (3.10) converge with probability 1 to a non-random limit on a compact set in  $\mathbb{C} \setminus \mathbb{R}$ .

Correspondingly, we prove in Sections 4 our second result

**II.** Let  $e_S^{(N)}(\Delta)$ ,  $g_S^{(N)}$ ,  $\nu$ ,  $H_S$  and  $\Sigma_S$  be defined by (3.4), (3.9) – (3.10), (2.1) – (2.2) and (2.5). Then:

(i) the limit  $e_S(\Delta)$  of (3.6) exists with probability 1, hence (3.2) and (3.7) hold with the same probability;

(ii) if  $f_S$  is the matrix Stieltjes transform of  $e_S$ , i.e.,

$$e_S(\Delta) = \lim_{\delta \rightarrow 0^+} \frac{1}{\pi} \int_{\Delta} \Im f_S(\lambda + i\delta) d\lambda \quad (3.14)$$

then  $f_S$  is a unique solution of the matrix equation

$$f_S(z) = \int (E + H_S - z - \Sigma_S f_S(z) \Sigma_S)^{-1} \nu(E) dE \quad (3.15)$$

in the class of matrix valued functions analytic for  $\Im z \neq 0$  and such that (cf. (3.12))

$$(f_S(z) - f_S^*(z))/(z - \bar{z}) > 0, \quad \Im z \neq 0, \quad (3.16)$$

and for any hermitian matrix  $A$  we write  $A > 0$  if  $A$  is positive definite.

**Remark.** The case  $n = 1$  of the above assertion corresponds to a particular case of the deformed Gaussian Ensembles of random matrix theory (see [22], Section 2, Theorem 2.2.1 in particular). The case  $n = 2$ ,  $H_S = s\sigma_z$ ,  $s > 0$ , and  $\Sigma_S = \sigma_z$ , where  $\sigma_x$  and  $\sigma_z$  are the corresponding Pauli matrices, was considered in [16, 17], while studying a random matrix model of quantum relaxation dynamics.

The obtained limiting reduced density matrix  $\rho_S$  (see (3.2) and (3.7)) is generally not of the Gibbs form (1.5). We thus need additional assumptions on the structure of reservoir in order to obtain the Gibbs distribution in the frameworks of our model. In our previous work [16], where the case  $n = 2$  was considered, it was assumed that the reservoir consists of a large number  $J \rightarrow \infty$  of practically non-interacting "macroscopically infinitesimal" but also sufficiently large parts (i.e., a kind of "coarse grained" structure of reservoir). This can be implemented by writing the density of states (2.2) of reservoir as the convolution of  $J$  copies of a certain density  $q$ :

$$\nu := \nu_J = q^{*J}. \quad (3.17)$$

The fact that we are going to consider the asymptotic regime  $J \rightarrow \infty$  after the limit  $N \rightarrow \infty$  can be interpreted as a formalization of the inequality determining our intermediate scale

$$n \ll J \ll N, \quad (3.18)$$

or, denoting  $n_q = N/J$  the parameter characterizing the "size" of infinitesimal parts, as the condition  $n_q \gg 1$ .

A simple example of the above is the Gaussian density

$$q(\varepsilon) = \frac{1}{(2\pi a^2)^{1/2}} \exp \left\{ -\frac{(\varepsilon - \varepsilon_0)^2}{2a^2} \right\}, \quad (3.19)$$



where  $\varepsilon_0$  is assumed to be of the order of magnitude of the characteristic energies of "macroscopically infinitesimal" parts of  $R$ , and  $a$  is their energy spread. In this case the density of states (3.17) of the reservoir is also Gaussian

$$\nu_J(E) = \frac{1}{(2\pi Ja^2)^{1/2}} \exp \left\{ -\frac{(E - J\varepsilon_0)^2}{2Ja^2} \right\}. \quad (3.20)$$

The formula makes explicit one more property of our "macroscopic" reservoir: its characteristic energies are of order of  $J\varepsilon_0$  with the spread  $(Ja^2)^{1/2}$ , while those of  $H_S$  and  $\Sigma_S$  are independent of  $J$ , hence, much smaller.

The Gaussian "ansatz" (3.19) – (3.20) was used in [16, 17]. In this paper we study the equilibrium properties of (2.5) with an arbitrary finite  $n$  and  $n \times n$  hermitian  $H_S$  and  $\Sigma_S$  in (2.5). As already noted, our results are valid for a wide and natural class of "macroscopically infinitesimal" part distributions  $q$  in (3.17), see Section 5. In particular, one can mention the densities

$$q_d = q_1^{*d}, \quad q_1(\varepsilon) = \left( \pi \sqrt{\varepsilon(\varepsilon_0 - \varepsilon)} \right)^{-1}, \quad 0 \leq \varepsilon \leq \varepsilon_0, \quad (3.21)$$

where  $q_1$  is the density of states of the one dimensional harmonic chain and  $q_d$  is the density of states of the  $d$ -dimensional harmonic cubic lattice. Different choices of  $q$  can model different types of reservoirs. For carrying out the proofs we shall require two conditions on  $q$ :

- (i)  $q$  decays superexponentially at  $-\infty$ ;
- (ii) there exists  $1 < a \leq 2$  such that

$$\int q^a(\varepsilon) d\varepsilon < \infty. \quad (3.22)$$

The conditions seem fairly natural. Indeed, condition (i) requires a rather "thin" if any spectrum of negative energies of large absolute value, thereby it is closely related to the stability of the reservoir. Furthermore, since (3.22) is valid by definition for  $a = 1$  (recall that  $q$  is a probability density), condition (ii) requires a certain regularity of the density of states, also usually assumed in statistical mechanics. Note also that if  $q$  is zero for large negative energies, then one can use the polynomial decay of the Fourier transform of  $q$  as another condition of its regularity.

Here are simple examples of the above. The first is the Gaussian density (3.19), for which the validity of the Gibbs distribution in the asymptotic regime (3.18) for  $n = 2$  was proved in [16]. Here  $a$  is any real number of  $(1, 2]$ . The second example is the exponential density  $q(\varepsilon) = e^{-\varepsilon/\varepsilon_0} \varepsilon_0^{-1}$ , where again  $a \in (1, 2]$ . The third example is the density of states (3.21) of the simple cubic  $d$ -dimensional crystal. Here  $a \in (1, 2)$  for  $d = 1$  and  $a \in (1, 2]$  for  $d \geq 2$ .

Viewing the density of states (3.17) as the partition function of the microcanonical ensemble of an ideal gas of  $J$  particles having each the partition function  $q$  one can introduce an analog of the entropy per particle

$$s(\varepsilon) := \lim_{J \rightarrow \infty} J^{-1} \log \nu_J(J\varepsilon). \quad (3.23)$$

The existence of the limit, its continuity and convexity can be proved by now standard argument of statistical mechanics (see e.g. [25, 20]).

One can also introduce an analog of the inverse temperature (cf. (1.9))

$$\beta(\varepsilon) = s'(\varepsilon). \quad (3.24)$$



The corresponding quantities for the canonical ensemble of "macroscopically infinitesimal" parts of our reservoir are the analogs of the partition function and the free energy per particle

$$Z_J(\beta) := \int e^{-\beta E} \nu_J(E) dE = \psi^J(\beta), \quad f(\beta) := -(\beta J)^{-1} \log Z_J(\beta) = -\beta^{-1} \log \psi(\beta), \quad (3.25)$$

where

$$\psi(\beta) = \int e^{-\beta \varepsilon} q(\varepsilon) d\varepsilon. \quad (3.26)$$

In writing the partition function (3.25) we took into account that the Laplace transform of the  $J$ -fold convolution  $\nu_J$  of (3.17) is  $J$ th power of the Laplace transform of  $q$ . Condition (i) above ensures that  $\psi$  is well defined for all non-negative  $\beta$ .

We have also from (3.23) and (3.25) – (3.26).

$$\beta f(\beta) = \min_{\varepsilon} \{\varepsilon \beta - s(\varepsilon)\} \quad (3.27)$$

and

$$s(\varepsilon) = \min_{\beta} h_{\varepsilon}(\beta), \quad h_{\varepsilon}(\beta) = \varepsilon \beta - f(\beta) \quad (3.28)$$

Note that  $\beta f$  and  $s$  are convex. The convexity of  $\beta f$  is immediate, since (3.25) and (3.26) imply

$$(\beta f)'' = - \left( \frac{\psi''}{\psi} - \frac{\psi'^2}{\psi^2} \right) = -d, \quad d := \int (\varepsilon - \varepsilon_0)^2 Q(\varepsilon) d\varepsilon > 0 \quad (3.29)$$

where

$$Q(\varepsilon) = e^{-\beta \varepsilon} q(\varepsilon) \left( \int e^{-\beta \varepsilon'} q(\varepsilon') d\varepsilon' \right)^{-1} \geq 0 \quad (3.30)$$

and  $\varepsilon_0$  is the first moment of  $Q$ . As for the convexity of  $s$ , it follows from the proof of (3.23) [25, 20].

We have also from (3.25) – (3.28)

$$h_{\varepsilon}(0) = 0, \quad h'_{\varepsilon}(0) = \varepsilon - \bar{\varepsilon}, \quad (3.31)$$

where

$$\bar{\varepsilon} = \int \varepsilon q(\varepsilon) d\varepsilon$$

is the mean energy of the macroscopically infinitesimal components of the reservoir. We will assume that

$$\varepsilon > \bar{\varepsilon}, \quad (3.32)$$

since then the convexity of  $h_{\varepsilon}$  and (3.31) imply that the minimum  $s(\varepsilon)$  of  $h_{\varepsilon}$  in (3.28) is positive, hence our temperature (3.24) is positive as well

$$\beta(\varepsilon) > 0, \quad (3.33)$$

although negative temperatures can also be considered in the frameworks of our model.

It turns out, however, that the above standard formulas of statistical mechanics of macroscopic systems are not accurate enough to obtain the Gibbs form (1.5) of the reduced density

matrix (3.14) – (3.15) of  $S$  in our model (3.17) of reservoir. This is because of the ”logarithmic accuracy” of the formulas, see e.g. (3.23), giving the large  $J$  leading term of  $\log \nu_J(J\varepsilon)$ , while we will need below the more accurate asymptotics formula:

$$\nu_J(J\varepsilon) = \mu_J(\varepsilon)(1 + o(1)), \quad J \rightarrow \infty, \quad \mu_J(\varepsilon) := (2\pi J s''(\varepsilon))^{-1/2} e^{-J s(\varepsilon)}. \quad (3.34)$$

This formula can be viewed as the Darwin-Fowler version of the equivalence of the micro-canonical and canonical ensembles in statistical mechanics (see e.g. [11]).

In fact, we do not need to know that  $\mu_J(\varepsilon)$  is the asymptotic (3.34) of  $\nu_J(J\varepsilon)$  to prove our third result below, i.e., the validity of the Gibbs form of the energy distribution of the system in our model (3.17) of the reservoir. Instead, we will just use  $\mu_J$  of (3.34) as an ansatz. Since, however, our proof is applicable to all finite  $n \geq 1$  and all  $n \times n$  hermitian  $H_S$  and  $\Sigma_S$  in (2.5) and since the case  $n = 1$ , and  $H_S = \Sigma_S = 0$  corresponds to the reservoir itself, thus, in this case  $\gamma_J(E) = \nu_J(E)$ , we obtain the asymptotic formula (3.34) as the simplest case  $n = 1$  of (3.36).

**III.** *Under the conditions of validity of Results I and II above and for the density of states of reservoir given by (3.17) with  $q$  satisfying conditions (i) and (ii), (5.2) and (3.22) in particular, we have:*

(i) *the reduced density matrix of (3.2) and (3.7) has a well defined limit as the width  $\delta$  of the energy shell  $\Delta$  of (1.7) tends to zero, i.e.,  $\Delta \rightarrow \{E\}$  and  $J$  is large enough:*

$$\hat{\rho}_J(E) := \lim_{\Delta \rightarrow \{E\}} \rho_S = \frac{\gamma_J(E)}{\text{Tr}_S \gamma_J(E)}, \quad \gamma_J(E) = \lim_{\Delta \rightarrow \{E\}} e_S(\Delta), \quad (3.35)$$

where  $e_S(\Delta)$  is given by (3.6);

(ii) *the limiting relation*

$$\lim_{J \rightarrow \infty} \gamma_J(J\varepsilon) / \mu_J(\varepsilon) = e^{-\beta H_S}, \quad (3.36)$$

with  $\mu_J$  given in (3.34), hence the Gibbs form (1.5) of the limit (in view of (3.35))

$$\lim_{J \rightarrow \infty} \hat{\rho}_J(J\varepsilon) = \frac{e^{-\beta H_S}}{\text{Tr}_S e^{-\beta H_S}}, \quad (3.37)$$

in which  $\beta := \beta(\varepsilon)$  is defined by (3.23) – (3.24) and (3.32) – (3.33).

The result is proved in Section 5.

**Remark.** The above result concerns the microcanonical energy distribution (1.6) – (1.8) of the composite system and the corresponding reduced density matrix (3.1). The case of the canonical distribution (1.2) of the whole system and the corresponding reduced density matrix, i.e., the passage from (1.4) to (1.5) can be readily obtained from (3.35) – (3.37). Indeed, we have from (1.2), the spectral theorem for  $H_{S \cup R}$ , (3.4) and (3.35) that with probability 1:

$$\lim_{N \rightarrow \infty} \frac{\text{Tr}_R e^{-\beta H_{S \cup R}}}{\text{Tr}_S \text{Tr}_R e^{-\beta H_{S \cup R}}} = \lim_{N \rightarrow \infty} \frac{\int e^{-\beta E} e_S(dE)}{\int e^{-\beta E} \text{Tr}_S e_S(dE)} = \frac{\int e^{-\beta E} \gamma_J(E) dE}{\int e^{-\beta E} \text{Tr}_S \gamma_J(E) dE}. \quad (3.38)$$

Using then (3.34), (3.36) and a simple saddle point argument for integrals on the r.h.s. of (3.38) yield

$$\lim_{J \rightarrow \infty} \frac{\int e^{-J(\beta\varepsilon - s(\varepsilon))} \gamma_J(J\varepsilon) \mu_J^{-1}(\varepsilon) d\varepsilon}{\int e^{-\beta E} \text{Tr}_S \gamma_J(J\varepsilon) \mu_J^{-1}(\varepsilon) d\varepsilon} = \frac{e^{-\beta H_S}}{\text{Tr}_S e^{-\beta H_S}}, \quad (3.39)$$

i.e., the Gibbs distribution again.

## 4 Selfaveraging Property and Limiting Reduced Density Matrix

In this section we prove Results 1 and II above. We note that the corresponding assertions as well as their proofs are generalizations of those for the deformed semicircle law (DSCL) of random matrix theory, see [22], Sections 2.2 and 18.3. Just as in the case of DSCL the passage from the positive definite matrix measures (positive measures in the case of DSCL) to their Stieltjes transforms (see (3.5) and (3.10)) reduces the proof of (3.1) – (3.2) to that of (3.15) – (3.16). The latter facts will follow if we prove that the variance of all the entries of (3.10) vanish fast enough as  $n \rightarrow \infty$  and that the expectation of (3.10) converges to a unique solution of (3.15) as  $n \rightarrow \infty$ .

We will use the Greek indices varying from 1 to  $n$  to label the states of the systems and the Latin indices varying from 1 to  $N$  to label the states of reservoir.

Then we can write  $n \times n$  and  $nN \times nN$  matrices of (3.10) and (3.8) as

$$g_S^{(N)}(z) = \{g_{\alpha\beta}(z)\}_{\alpha,\beta=1}^n, \quad \{G_{\alpha j, \beta j}(z)\}_{\alpha,\beta=1; j,k=1}^{n,N} \quad (4.1)$$

and

$$g_{\alpha\beta}(z) = \frac{1}{N} \sum_{j=1}^N G_{\alpha j, \beta j}(z). \quad (4.2)$$

We will prove now the bound

$$\mathbf{Var}\{g_{\alpha\beta}(z) := \mathbf{E}\{|g_{\alpha\beta}(z)|^2\} - |\mathbf{E}\{g_{\alpha\beta}(z)\}|^2 \leq C_S(z)/N^2, \quad (4.3)$$

and we write here and below  $C_S(z)$  for quantities which do not depend on  $N$  and are finite for  $\Im z \neq 0$ .

To this end we view every  $g_{\alpha\beta}$  as a function of the Gaussian random variables  $\{W_{jk}\}_{j,k=1}^N$  of (2.3) – (2.4) and use the Poincaré inequality (see [22], Proposition 2.1.6), yielding

$$\mathbf{Var}\{g_{\alpha\beta}(z)\} \leq \sum_{j,k=1}^N \mathbf{E} \left\{ \left| \frac{\partial g_{\alpha\beta}}{\partial W_{jk}} \right|^2 \right\}. \quad (4.4)$$

The derivatives on the right can be found by using the resolvent identity for  $G$

$$\begin{aligned} \frac{\partial g_{\alpha\beta}(z)}{\partial W_{jk}} &= -\frac{1}{N^{3/2}} \sum_{\gamma, \delta=1}^n (G_{\delta\beta}(z) G_{\alpha\gamma}(z))_{jk} \Sigma_{\gamma\delta} \\ &\leq \left( \frac{1}{N^3} \sum_{\gamma, \delta=1}^n |(G_{\delta\beta}(z) G_{\alpha\gamma}(z))_{jk}|^2 \sum_{\gamma, \delta=1}^n |\Sigma_{\gamma\delta}|^2 \right)^{1/2}, \end{aligned}$$

where  $G_{\alpha\beta}$  denotes the  $N \times N$  matrix

$$G_{\alpha\beta}(z) = \{G_{\alpha j, \beta k}(z)\}_{j,k=1}^N.$$

The  $n \times n$  matrix  $\Sigma_S = \{\Sigma_{\alpha\beta}\}_{\alpha,\beta=1}^n$  is defined by (2.5) and the second line results from Schwarz's inequality. Taking into account the bounds

$$\|G(z)\| \leq |\Im z|^{-1}, \quad |G_{\alpha j, \beta k}(z)| \leq \|G(z)\| \leq |\Im z|^{-1} \quad (4.5)$$

valid for the resolvent of any hermitian matrix, it is easy to find an analogous bound for the matrix  $G_{\alpha\beta}(z)$ :

$$||G_{\alpha\beta}(z)|| \leq |\Im z|^{-1}. \quad (4.6)$$

The above relations imply for the r.h.s. of (4.4)

$$\mathbf{Var}\{g_{\alpha\beta}(z)\} \leq \frac{1}{N^3} \text{Tr}_S \Sigma_S^2 \sum_{\gamma, \delta=1}^n \text{Tr}_R (G_{\delta\beta} G_{\alpha\gamma} G_{\alpha\gamma}^* G_{\delta\beta}^*)$$

and using now (4.6) and the inequality  $|\text{Tr} A| \leq N ||A||$  valid for any  $N \times N$  matrix, we obtain

$$\mathbf{Var}\{g_{\alpha\beta}(z)\} \leq \frac{n^2}{N^2 |\Im z|^4} \text{Tr}_S \Sigma_S \Sigma_S^*,$$

i.e., the bound (4.3) with  $C_S(z) = n^2 \text{Tr}_S \Sigma_S^2 / |\Im z|^4$ .

Denote

$$f_S^{(N)}(z) = \{f_{\alpha\beta}(z)\}_{\alpha, \beta=1}^n, \quad f_{\alpha\beta}(z) = \mathbf{E}\{g_{\alpha\beta}(z)\}. \quad (4.7)$$

We will now prove that for any compact set  $K \subset \mathbb{C} \setminus \mathbb{R}$  there is a subsequence  $\{f_S^{(N_i)}\}_{N_i}$  of analytic matrix functions, which converges on  $K$  as  $N_i \rightarrow \infty$  to a solution of (3.15). Repeating almost literally the argument, which leads to Req. (2.2.8) of [22], we obtain

$$\mathbf{E}\{G(z)\} = \left( (H_S - \Sigma_S f_S^{(N)}(z) \Sigma_S) \otimes \mathbf{1}_R + \mathbf{1}_S \otimes H_R - z \mathbf{1}_{S \cup R} \right)^{-1} + R_N(z), \quad (4.8)$$

where the  $nN \times nN$  matrix  $R_N(z)$  admits the bound

$$||R_N(z)|| \leq C_S(z)/N^2. \quad (4.9)$$

Applying to (4.8) – (4.9) the formula

$$\text{Tr}_R \varphi(A_S \otimes \mathbf{1}_R + \mathbf{1}_S \otimes B_R) = \sum_{l=1}^N \varphi(A_S + b_l),$$

valid for any function  $\varphi$ , any  $n \times n$  matrix  $A_S$  and any  $N \times N$  hermitian matrix  $B_R$  with eigenvalues  $\{b_l\}_{l=1}^N$  and using the definition (2.1) of the density of states of the reservoir, we obtain for the expectation (4.7)

$$f_S^{(N)}(z) = \int \frac{\nu_N(E) dE}{E + H_S - z - \Sigma_S f_S^{(N)}(z) \Sigma_S} + r_N(z), \quad (4.10)$$

where

$$||r_N(z)|| \leq C_S(z)/N^2 |\Im z|^3.$$

It follows from (4.7) and (4.5) that

$$||f_S^{(N)}(z)|| \leq |\Im z|^{-1}.$$

Hence, for any compact set  $K \subset \mathbb{C} \setminus \mathbb{R}$  there exists a subsequence  $\{f_S^{(N_i)}\}_{N_i}$  of bounded analytic matrix functions, which converges uniformly on  $K$  to an analytic matrix function  $f$

and passing to the limit  $N \rightarrow \infty$  in (4.10) we obtain (3.15) for  $z \in K$ . The validity of the equation for any  $z \in \mathbb{C} \setminus \mathbb{R}$  follows from the analyticity of  $f$  in  $\mathbb{C} \setminus \mathbb{R}$ .

Let us show that (3.15) is uniquely solvable in the class of analytic matrix functions (3.16). Choose

$$K \subset \{z \in \mathbb{C} : |\Im z| > \|\Sigma_S\|^2\} \quad (4.11)$$

and assume that there are two different solutions  $f'_S$  and  $f''_S$ , i.e.,  $\max_{z \in K} \|f'_S(z) - f''_S(z)\| > 0$ . It follows then from (3.15) and (3.16) that

$$1 \leq \|\Sigma_S\|^2 \int dE \nu(E) \|(E - H_S - z - \Sigma_S f'_S \Sigma_S)^{-1}\| \cdot \|(E - H_S - z - \Sigma_S f''_S \Sigma_S)^{-1}\|. \quad (4.12)$$

In addition, (3.16) implies the bounds  $\|(E - H_S - z - \Sigma_S f'_S \Sigma_S)^{-1}\| \leq |\Im z|^{-1}$  and  $\|(E - H_S - z - \Sigma_S f''_S \Sigma_S)^{-1}\| \leq |\Im z|^{-1}$  (see (5.4) for details). This and (4.12) yield the inequality  $1 \leq \|\Sigma_S\|^2 / |\Im z|^2$ , which contradicts (4.11).

The unique solvability of (3.15) – (3.16) implies that the whole sequence  $\{f_S^{(N)}\}_N$  converges to the limit  $f$ , a unique solution of (3.15) – (3.16) everywhere in  $\mathbb{C} \setminus \mathbb{R}$ .

It remains to prove that this fact together with (4.3) imply the convergence with probability 1 of the sequence  $\{g_S^{(N)}\}_N$  of random analytic matrix functions to the solution of (3.15) – (3.16).

It follows from the Tchebyshev inequality and (4.3) that for any  $\delta > 0$  and any non-real  $z$

$$\mathbf{P}\{\|f_S^{(N)}(z) - g_S^{(N)}(z)\| > \delta\} \leq \mathbf{Var}\{g_S^{(N)}(z)\} / \delta^2 \leq C_S(z) / \delta^2 |\Im z|^4 N^2.$$

Hence, for any non-real  $z$  the series

$$\sum_{N=1}^{\infty} \mathbf{P}\{\|f_S^{(N)}(z) - g_S^{(N)}(z)\| > \delta\}$$

converges for any  $\delta > 0$ , and by the Borel-Cantelli lemma and the convergence of  $\{f_S^{(N)}\}_N$  to  $f_S$  we have with probability 1 for any non-real  $z$

$$\lim_{N \rightarrow \infty} g_S^{(N)}(z) = f_S(z). \quad (4.13)$$

Let us show that  $g_S^{(N)}$  converges to  $f_S$  uniformly on any compact set of  $\mathbb{C} \setminus \mathbb{R}$  with probability 1. Because of the uniqueness of analytic continuation it suffices to prove that with the same probability the limiting relation  $\lim_{N \rightarrow \infty} g_S^{(N)}(z_j) = f_S(z_j)$  is valid for all points of an infinite sequence  $\{z_j\}_{j \geq 1}$ ,  $\Im z_j \geq \eta_0 > 0$ , possessing a finite accumulation point. Denote  $\Omega(z)$  the set of realizations, where (4.13) is valid. According to the above  $\mathbf{P}\{\Omega(z_j)\} = 1$ ,  $\forall j$ . Hence,

$$\mathbf{P}\left\{\bigcap_{j \geq 1} \Omega(z_j)\right\} = 1.$$

This proves the uniform convergence of  $g_S^{(N)}$  to  $f_S$  on any compact set of  $\mathbb{C} \setminus \mathbb{R}$  with probability 1.

It is known that the one-to-one correspondence between the non negative measures and the Stieltjes transforms is continuous in the topology of vague convergence of measures and the uniform convergence on a compact set of  $\mathbb{C} \setminus \mathbb{R}$  [22], Proposition 2.1.2. An analogous assertion can be easily proved for the positive definite matrix measures and their Stieltjes transforms. By using this fact and the uniform convergence with probability 1 of  $g_S^{(N)}$  to  $f_S$  proved above, we obtain (3.6).

## 5 Gibbs Distribution

We will prove here our Result III, i.e., relation (3.37), by using Result II (see (3.15) – (3.16)) and our model (3.17) of the reservoir. Recall that we assume that the density  $q$  of (3.17) satisfies the conditions (i) and (ii) of Section 3, (5.2) – (3.22) in particular.

We will now use the matrix identity (see formulas (5.4) for its justification)

$$-\frac{1}{\zeta} = i \int_0^\infty e^{it\zeta} dt, \quad \Im \zeta > 0,$$

with  $\zeta = z + \Sigma_S f_S(z) \Sigma_S - E - H_S$  to write the basic equation (3.15) with  $\Im z > 0$  as

$$f_S(z) = \int_0^\infty i e^{it(-\tilde{H}+z)} \varphi^J(t) dt, \quad \tilde{H} = H_S - \Sigma_S f_S(z) \Sigma_S, \quad (5.1)$$

where

$$\varphi(t) = \int e^{-i\varepsilon t} q(\varepsilon) d\varepsilon, \quad \varphi(0) = 1, \quad |\varphi(t)| \leq 1, \quad (5.2)$$

and we took into account that the Fourier transform of the convolution (3.17) is  $\varphi^J$ . Our proofs in this section are based on the above integral representation of  $f_S$ .

Let us prove first that the limit

$$f_S(E) := \lim_{\delta \rightarrow 0^+} f_S(E + i\delta) \quad (5.3)$$

exists and is bounded for all real  $E$  if  $J$  is large enough. To this end we first use the bounds

$$\|(M - \zeta)^{-1}\| \leq (\Im \zeta)^{-1}, \quad \Im \zeta > 0; \quad \|e^{-itM}\| \leq 1, \quad t > 0, \quad (5.4)$$

valid for any complex matrix such that  $\Im M > 0$ . Indeed, in this case  $\Re(-iM) \geq 0$ . Matrices with this property are known as accretive and for them the corresponding bounds  $\|(-iM + \zeta)^{-1}\| \leq (\Re \zeta)^{-1}$ ,  $\Re \zeta > 0$ ;  $\|e^{-tM}\| \leq 1$ ,  $t > 0$  are valid (see e.g. [14], Section IX.1.6). Noting that  $\Im f_S(z) > 0$ ,  $\Im z > 0$  implies  $\Im \Sigma f_S(z) \Sigma > 0$  for  $\Im z > 0$  and using the second bound of (5.4), we obtain from (5.1)

$$\|f_S(z)\| \leq \int_0^\infty |\varphi(t)|^J dt, \quad \Im z > 0. \quad (5.5)$$

It follows then from (3.22) and the Hausdorff-Young inequality that

$$\int |\varphi(t)|^{J_0} dt \leq \left( \int q^a(\varepsilon) d\varepsilon \right)^{J_0/a} < \infty, \quad J_0 = a/(a-1) \quad (5.6)$$

and we obtain from (5.2), (5.5) and (5.6)

$$\|f_S(z)\| \leq \int_0^\infty |\varphi(t)|^{J_0} dt < \infty, \quad J \geq J_0. \quad (5.7)$$

It is important that the r.h.s. of the above bound is independent of  $z$  for  $\Im z \geq 0$  and  $J \geq J_0$ , thus the same bound holds for  $f_S(E)$  of (5.3) with any real  $E$  and uniformly in  $J \geq J_0$ .

We will use now a stronger version of the inversion formula (3.13) corresponding to measures with a bounded density  $m'$ :

$$m'(\lambda) := \lim_{\Delta \rightarrow \{\lambda\}} m(\Delta)/|\Delta| = \lim_{\delta \rightarrow 0^+} \pi^{-1} \Im s(\lambda + i\delta). \quad (5.8)$$

Since the same formula holds also in the matrix case, we will divide the numerator and denominator of (3.7) by  $|\Delta|$  and pass to the limit  $\Delta \rightarrow \{E\}$  to obtain (3.35).

In fact, an elaboration of the above argument allows us to prove the relation

$$\lim_{J \rightarrow \infty} \|f_S(J\varepsilon)\| = 0, \quad (5.9)$$

valid for the limit (5.3) with any  $\varepsilon$  varying over a finite interval which will be used below.

Indeed, it is shown in the Appendix that if  $\varphi$  satisfies (3.22), then for any  $t_0 > 0$

$$\max_{t \geq t_0 > 0} |\varphi(t)| = \varphi_0 < 1. \quad (5.10)$$

Choose a  $\delta > 0$  and write the integral in (5.5) as the sum of integrals over  $(0, \delta)$  and  $(\delta, \infty)$ . Then it follows from (5.2), (3.22) and (5.6) that the first integral is bounded by  $\delta$  and the second is bounded by

$$\varphi_0^{J-J_0} \int_{\delta}^{\infty} |\varphi(t)|^{J_0} dt \leq C \varphi_0^J, \quad J \geq J_0, \quad (5.11)$$

where  $C$  is independent of  $J$  and  $\delta$ . Thus, passing to the limit  $J \rightarrow \infty$  and then  $\delta \rightarrow 0$ , we obtain (5.9).

In fact, a bit more careful calculation yields that the r.h.s. of (5.9) is of the order  $J^{-1/2}$  if the first moment of  $q$  is zero, and  $J^{-1}$  if the first moment is not zero. This can be seen also in examples (3.19) and (3.21).

Introduce the real and imaginary part of  $f_S(J\varepsilon)$  (note that  $f_S(J\varepsilon)$  is well defined in view of (5.5)):

$$f_S(J\varepsilon) = \mathcal{R} + i\mathcal{I} \quad (5.12)$$

and take into account that according to our assumptions the function  $\varphi$  of (5.2) can be analytically continued into the lower half-plane in  $t$ . This allows us to write the integral for  $\mathcal{I}$ , which determines  $\gamma_J$  of (3.35) according to (5.8), as the sum

$$\mathcal{I} := \Im f_S(J\varepsilon) = I_1 + I_2 \quad (5.13)$$

of integrals over  $(0, -i\beta(\varepsilon))$  and  $(-i\tau\beta(\varepsilon), -i\beta(\varepsilon) + \infty)$ , where  $\beta(\varepsilon) > 0$  is the point where the function  $h_\varepsilon$  of (3.28) achieves its minimum, see (3.31) – (3.33):

Changing  $t$  to  $-i\tau$  in  $I_1$  we obtain

$$I_1 = \int_0^\beta e^{Jh_\varepsilon(\tau)} \mathcal{S}(\tau) d\tau, \quad (5.14)$$

where we write

$$e^{-\tau\tilde{H}} = \mathcal{C}(\tau) + i\mathcal{S}(\tau). \quad (5.15)$$

for  $\tilde{H}$  of (5.1) and  $\beta$  instead of  $\beta(\varepsilon)$ . Write also

$$\tilde{H} = R + iI, \quad R = H_S - \Sigma_S \mathcal{R} \Sigma_S, \quad I = \Sigma_S \mathcal{I} \Sigma_S \geq 0. \quad (5.16)$$



If  $\tilde{H}$  were a complex number but not a matrix (e.g., if  $n = 1$ ), then

$$\mathcal{S}(\tau) = -e^{-\tau R} \sin \tau I = -e^{-\tau R} I c(\tau), \quad |c(\tau)| \leq \tau \quad (5.17)$$

and

$$I_1 = - \int_0^\beta e^{J h_\varepsilon(\tau)} e^{-\tau R} I c(\tau) d\tau. \quad (5.18)$$

This and the inequality (see (3.28) – (3.32))

$$h(\tau) \geq (\varepsilon - \bar{\varepsilon})\tau, \quad (5.19)$$

would imply the bound

$$I_1 \leq e^{\beta|R|} |I| \int_0^\beta \tau e^{-J|\varepsilon - \bar{\varepsilon}|\tau} d\tau \leq e^{\beta|R|} |\Sigma_S|^2 \mathcal{I} / ((\varepsilon - \bar{\varepsilon})J)^2. \quad (5.20)$$

Comparing this with (5.13), we would conclude that the contribution of  $I_1$  in  $\mathcal{I}$  of (5.13) is negligible as  $J \rightarrow \infty$ , i.e.,

$$\mathcal{I} = I_2 (1 + o(1)), \quad J \rightarrow \infty. \quad (5.21)$$

We are going now to find a matrix analog of the above asymptotic relation.

We note first that according to Appendix the matrix analog of (5.17) is

$$\mathcal{S}(\tau) = - \int_0^\tau e^{-(\tau-s)R} I \mathcal{C}(s) ds, \quad \|\mathcal{C}(s)\| \leq s e^{s\|R\|} \cosh(s\|I\|), \quad (5.22)$$

hence the matrix analog of (5.18) is in view of (5.16)

$$\begin{aligned} I_1 &= - \int_0^\beta e^{J h_\varepsilon(\tau)} d\tau \int_0^\tau e^{-(\tau-s)R} I \mathcal{C}(s) ds \\ &= \int_0^\beta e^{J h_\varepsilon(\tau)} d\tau \int_0^\tau e^{-(\tau-s)R} \Sigma_S \mathcal{I} \Sigma_S \mathcal{C}(s) ds. \end{aligned}$$

It is convenient to view the r.h.s. of the above formula as the result of application to the  $n \times n$  matrix  $\mathcal{I}$  of the linear operator  $\mathbf{A}$  acting in the space  $\mathcal{M}_n(\mathbb{C})$  of complex  $n \times n$  matrices, i.e., to write the formula as

$$I_1 = \mathbf{A}(\mathcal{I}). \quad (5.23)$$

Then we have from (5.22) just as in the scalar case (5.17) – (5.20):

$$\|I_1\| = \|\mathbf{A}(\mathcal{I})\| \leq e^{\beta\|R\|} \|\Sigma_S\|^2 \|\mathcal{I}\| \cosh(\beta\|\Sigma_S\|^2 \|\mathcal{I}\|) / ((\varepsilon - \bar{\varepsilon})J)^2.$$

By using the bounds  $\|R\| \leq 2\|H\|$  and  $\|\mathcal{I}\| \leq 1$ , which following from (5.9), we conclude that the matrix analog of bound (5.21) is also valid, i.e., the contribution of  $I_1$  to the r.h.s. of (5.13) is negligible as  $J \rightarrow \infty$  in the matrix case as well. This fact can be expressed via the operator  $\mathbf{A}$ :

$$\|\mathbf{A}\| \leq C_S / ((\varepsilon - \bar{\varepsilon})J)^2. \quad (5.24)$$

Consider now  $I_2$  of (5.13). Changing the variable to  $t = -i\tau_0 + \sigma$  and using  $\mu_J$  of (3.34), we obtain

$$I_2 / \mu_J(\varepsilon) = \sqrt{2\pi s''(\varepsilon)J} \Re e^{-\beta\tilde{H}} \int_0^\infty e^{-i\sigma\tilde{H}} e^{J\chi(\sigma)} d\sigma, \quad (5.25)$$

where

$$\chi(\sigma) = i\varepsilon\tau + \log \psi_\beta(\sigma), \quad \psi_\beta(\sigma) = \psi(\beta + i\sigma)/\psi(\beta). \quad (5.26)$$

We write the integral on the right of (5.25) as the sum of integrals over  $(0, \sigma_0)$  and  $(\sigma_0, \infty)$ , where  $\sigma_0$  is small enough:

$$I_2 = I_{21} + I_{22}. \quad (5.27)$$

In the first integral we use the expansion  $\chi(\sigma) = 2^{-1}s''(\varepsilon)\sigma^2(1 + o(1))$ ,  $\sigma \rightarrow 0$  of (5.26) yielding

$$I_{21} = \pi \Re e^{-\beta \tilde{H}}(1 + o(1)), \quad J \rightarrow \infty,$$

and then (5.9) and (5.16) imply

$$I_{21} = \pi e^{-\beta Hs}(1 + o(1)), \quad J \rightarrow \infty. \quad (5.28)$$

To deal with the second term

$$I_{22} = \Re \int_{\sigma_0}^{\infty} e^{-(\tau_0 + i\sigma)\tilde{H}} e^{J\chi(\sigma)} d\sigma$$

of (5.27) we use first (5.9) and (5.16) to obtain the bound  $\|e^{-(\beta + i\sigma)\tilde{H}}\| \leq K(\varepsilon)$ , where  $K(\varepsilon)$  is independent of  $J$ . This and (5.25) – (5.26) yield

$$\|I_{22}/\mu_J(\varepsilon)\| \leq \sqrt{2\pi s''(\varepsilon)JK(\varepsilon)} \int_{\delta}^{\infty} |\psi_\varepsilon(\sigma)|^J d\sigma, \quad (5.29)$$

where

$$\psi_\beta(\sigma) = \psi(\beta + i\sigma)/\psi(\beta).$$

To estimate the integral on the right of (5.29) we will follow the scheme of proof of (5.11), which was based on the bounds (5.10) and (3.22). Thus, we have to prove the analogs of these bounds for  $\psi_\beta$ . To this end we use (3.26) and (3.30) to write

$$\psi_\beta(\sigma) = \int e^{-i\varepsilon\sigma} Q(\varepsilon) d\varepsilon. \quad (5.30)$$

We conclude that the  $\psi_\beta$  is the Fourier transform of a non-negative function of unit integral, thus it satisfies (5.2), i.e.,  $|\psi_\beta(\sigma)| \leq 1$ ,  $\psi_\beta(0) = 1$ . It follows then from Appendix that in this case we have an analog of (5.10) as well:

$$\max_{\sigma \geq \sigma_0 > 0} |\psi_\beta(\sigma)| := \kappa_1 < 1. \quad (5.31)$$

Furthermore, according to (5.2) and (5.30), an analog of (3.22) for  $\psi_\beta$  is

$$\int Q^a(\varepsilon) d\varepsilon = \int e^{-a\beta\varepsilon} q^a(\varepsilon) d\varepsilon < \infty.$$

It is easy to see this is indeed true in view of conditions (i) and (ii) of Section 3, (3.22) in particular.

This, (5.27) and (5.28) yield  $I_2 = \pi e^{-\beta Hs}(1 + o(1))$ ,  $J \rightarrow \infty$ . Now use (5.23) to write (5.13) as  $\mathcal{I} = \mathbf{A}(\mathcal{I}) + I_2$ . Then (5.24) and (5.25) imply

$$\mathcal{I}/\mu_J(\varepsilon) = (\mathbf{1} - \mathbf{A})^{-1} I_2/\mu_J(\varepsilon) = \pi e^{-\beta Hs}(1 + o(1)), \quad J \rightarrow \infty.$$

This and (5.8) yield (3.36). It remains then to divide the numerator and the denominator in (3.35) by  $\mu_j$  and to use (3.36) to obtain (3.37).

**Acknowledgement.** We thank Chris Jarzynski for helpful comments. We also thank the Institute for Advanced Study in Princeton N.J. for hospitality when the part this work was done. Supported in part by NSF Grant DMR-1104501.

# Appendix

(i) *Proof of (5.10).* We prove here that if  $q$  is non-negative, continuous and of unit integral, then we have (5.10) in addition to (5.2).

Indeed, the equality  $\varphi(t') = 0$  for some  $t' \neq 0$  implies the equality

$$\int (1 - \cos \varepsilon t') q(\varepsilon) d\varepsilon = 0,$$

which is impossible for any non-negative and non-zero  $q$  satisfying (3.22). This implies (5.10).

(ii) *Proof of (5.22).* The Duhamel formula

$$e^{A+B} = e^A + \int_0^1 e^{(1-s)A} B e^{s(A+B)} ds,$$

valid for any two (generally non-commuting) matrices, yields the formulas for the terms of the r.h.s. of (5.15)

$$\mathcal{S}(\tau) = - \int_0^\tau e^{-(\tau-s)R} I C(s) ds,$$

and

$$C(\tau) = e^{-\tau R} + \int_0^\tau e^{-(\tau-s)R} I \mathcal{S}(s) ds$$

hence

$$\mathcal{C}(s) = e^{-sR} + \sum_{l=1}^{\infty} (-1)^l \int_0^s e^{(s-s_1)R} I ds_1 \int_0^{s_1} e^{-(s_1-s_2)R} I ds_2 \dots \int_0^{s_{2l-1}} e^{-(s_{2l-1}-s_{2l})R} I e^{-s_{2l}R} ds_{2l}.$$

These formulas combined with the standard upper bounds for the terms of the series lead to (5.22).

In the commutative case the formulas are

$$e^{-\tau R} \sin \tau I = e^{-\tau R} I c(\tau), \quad c(\tau) = \int_0^\tau \cos s I ds,$$

i.e., coincide with (5.17)). The non-commutative analog (5.22) of (5.17) is more rough, since we do not take into account the alternating signs in the above series, hence the effects of strong cancelations of its terms.

## References

- [1] Akemann G., Baik J., Di Francesco P. (Eds.) *The Oxford Handbook of Random Matrix Theory*, Oxford University Press, Oxford 2011.
- [2] Attal S., Joye A., Pillet C.-A. (Eds.) *Open Quantum Systems, I – III*, Lecture Notes in Mathematics, **1880 – 1882**, Springer, Berlin Heidelberg, 2006.
- [3] Bach, V., Fröhlich, J., Sigal, I.M. Return to equilibrium, *J. Math. Phys.* **41** (2000) 3985 – 4060.

- [4] Bogolyubov, N. *O Nekotoryh Statisticheskikh Metodah v Matematicheskoi Fizike*. (Russian) [*On Some Statistical Methods in Mathematical Physics*] Akademiya Nauk Ukrainskoi SSR, 1945. For modern version see Bogolubov, N. N., Bogolubov, N. N., Jr. *Introduction to quantum statistical mechanics*. World Scientific Publishing Co., Hackensack, NJ, 2010.
- [5] De Roeck W., Kupiainen A. "Return to Equilibrium" for Weakly Coupled Quantum Systems: A Simple Polymer Expansion, *Commun. Math. Phys.* **305** (2011) 797–826.
- [6] Gemmer J., Michel M., Mahler G. *Quantum Thermodynamics*, Springer Berlin 2005.
- [7] Goldstein S., Lebowitz J. L., Tumulka R., Zanghi N., Canonical typicality, *Phys. Rev. Lett.* **96** (2006) 050403.
- [8] Goldstein S., Lebowitz J. L., Mastrodonato C., Tumulka R., Zanghi N. On the approach to thermal equilibrium of macroscopic quantum systems, *Phys. Rev. E* **81** (2010) 011109.
- [9] Guhr T., Moeller-Groeling A., Weidenmueller H. A. Random-matrix theories in quantum physics: common concepts, *Phys. Rep.* 299 (1998) 189–425.
- [10] Huang K. *Lectures on Statistical Physics and Protein Folding*, World Sci., Singapore, 2005.
- [11] Huang K. *Statistical Mechanics*, Wiley, N.Y., 1987.
- [12] Jarzynski C. Nonequilibrium work theorem for a system strongly coupled to a thermal environment, *J. Stat. Mech.: Theor. Exp.* (2004) P00000.
- [13] Jarzynski C. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, *Annual Rev. Condens. Matter Phys.* **2** (2011) 329–351.
- [14] Kato T., *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1976.
- [15] Landau L., Lifshitz E. *Statistical Physics*, Pergamon, Oxford, 1980.
- [16] Lebowitz J. L., Pastur L. A random matrix model of relaxation, *J. Phys. A: Math. Gen.* **37** (2004) 1517–1534.
- [17] Lebowitz J. L., Lytova A., Pastur L. On a random matrix model of quantum relaxation, In: *Contemporary Mathematics, "Adventures in Mathematical Physics"*, vol. 447, F. Germinet, P. Hislop (Eds.), AMS, Providence, 2007, pp.199 – 218.
- [18] Lifshitz I. M., Gredeskul S. A., Pastur L. A., *Introduction to the Theory of Disordered Systems*, Wiley, N.-Y., 1988.
- [19] Linden N., Popescu S., Skrzypczyk P. How small can thermal machines be? The smallest possible refrigerator, *Phys. Rev. Lett.* **105** (2010) 130401, arXiv:0908.2076.
- [20] Minlos R. A., Povzner A. Ya. On the thermodynamic limit for entropy, *Proceedings of Moscow Mathematical Society*, **17** (1967) 243 – 272 (Russian).

- [21] Nandkishore R., Huse D.A. Many-body localization and thermalization in quantum statistical mechanics, arXiv:1404.0686.
- [22] Pastur L., Shcherbina M. *Eigenvalue Distribution of Large Random Matrices* AMS, Providence, RI, 2011.
- [23] Popescu S., Short A. J., Winter A. Entanglement and the foundations of statistical mechanics, *Nature Phys.* **2** (2006) 754 .(Full version e-print arXiv:quant-ph/0511225).
- [24] Ritort F. Nonequilibrium fluctuations in small systems: From physics to biology, *Adv. Chem. Phys.* **137** (2008) 31–123.
- [25] Ruelle D. *Statistical Mechanics*, Benjamin, N. Y., 1969.
- [26] Seifert U. Stochastic thermodynamics, fluctuation theorems, and molecular machines, *Rep. Prog. Phys.* **75** (2012) 126001.